Modeling and Forecasting Malaria in Tripura, INDIA using NOAA/AVHRR-Based Vegetation Health Indices

Mohammad Nizamuddin*[1], Felix Kogan[2], Ramesh Dhiman[3], Wei Guo[4], Leonid Roytman[1]

[1]NOAA-CREST, City College, New York, NY, USA

[2]NOAA/NESDIS/STAR, Camp Springs, MD, USA

[3]National Institute of Malaria Research (ICMR), New Delhi, India

[4]IMSG Inc, Camp Springs, MD, USA

*[1]mnizamuddin@ccny.cuny.edu; [2]Felix.Kogan@noaa.gov;[3]tdhimanrc@icmr.org; [4]Wei.Guo@NOAA.gov; [1]roytman@ccny.cuny.edu;

Abstract

Improved forecasting, prevention and control of epidemics are the key technical elements for malaria eradication program. The objective is to use NOAA/AVHRR environmental satellite data to produce weather seasonal forecasts for using as a proxy for predicting malaria epidemics in Tripura state, India which has one of the highest endemic of malaria cases in the country. An algorithm has been reported that uses Vegetation Health (VH) Indices (Vegetation Condition Index (VCI) and Temperature Condition Index (TCI)) computed from Advance Very High Resolution Radiometer (AVHRR) data flown on NOAA afternoon polar orbiting satellite. A significant relationship between satellite data and annual malaria incidences is found at least three months before the major malaria transmission period. Principal component regression (PCR) method was used to develop a model to predict malaria as a function of the TCI. The simulated results were compared with observed malaria statistics showing that the error of estimations of malaria is insignificant. Optical remote sensing therefore is a valuable tool to estimate malaria well in advance so that preventive measures can be taken.

Keywords

AVHRR; VCI; TCI; Malaria; Principal Component Regression

Introduction

As the 1.4 billion people living in 11 countries of the Southeastern Asia, around 1.2 billion are exposed to the risk of malaria (Shiv Lal et. el. 2010) . Every year, nearly 2.5 million of malaria cases are reported in these countries and 75-85% are reported in India. There are three major areas in India, where malaria infections exceed 10 persons per 1000 inhabitants per year (http://www.nvbdcp.gov.in). Tripura state is located in the eastern part of India, bordered by Bangladesh. Although this state is small in size (0.31% of the entire India), there is a very important consideration for its selection in this study: In addition to being a huge malaria burden in Tripura, Plasmodium (P) falciparum, which causes malaria has become resistant to drugs, especially in the southern part. This creates a big problem for the 3.2 million inhabitants in Tripura and also tourists and border patrol troops. Therefore, an early detection and monitoring of malaria in Tripura is warranted for protection of the population.

Since malaria is strongly weather dependent, epidemics occurrence, timing, area and intensity can be monitored by meteorological (precipitation, temperature, humidity) data (Thomson et. el. 2006). Unfortunately, for the 10,491 km^2 of Tripura area, there are only five weather stations, which is not sufficient for effective malaria monitoring. Therefore, this study investigates the application of operational satellites for the indicated purpose. Previous studies in the neighboring Bangladesh, which used the most advanced vegetation health (VH) methodology, clearly showed the utility of satellite based VH for early malaria detection and monitoring the number of people diagnosed with the disease (Rahman et al.2011, Rahman et el.2010). Therefore, we attempted to use VH indices developed from the Advanced Very High Resolution Radiometer (AVHRR) flown on NOAA-series of operational polar-orbiting satellites to monitor malaria epidemics in Tripura. Specifically, an attempt was made to develop malaria incidence-VH models using the Principal component analysis (PCA) and apply the models to early detection of malaria start, prediction of its dynamics and resultant malaria cases.

Study Area

Tripura state is located in the northeastern India (22°56'N to 24°32'N and 90°09'E to 92°10'E). The state is bordered with Bangladesh in the north, south, and west and the India's Assam and Mizoram states in the east. The area is accessible to the rest of India through the Karimganj district of Assam state and Aizawl district of Mizoram state in the east. Tripura is divided into four districts with the majority of the population (72%) living in the southern and western districts (*Wikipedia, 2010)*. Tripura is a landlocked hilly state with altitudes varying from 15 to 940 m above the sea level but the majority of population reside in the low land (plains). Besides, the area has many rivers among which the biggest one is the Manu River.

Tripura is a Malaria prone zone and three out of its four districts have been declared as anti malaria drug resistant. Every year, thousands of local people and also border guard, are falling prey to malaria in Tripura (http://www.indigenousherald.com/health/85-border-troops-tame-malaria-with-face-mask.html). In the recent years, inhabitants of the interior Tripura have stopped recovering from the application of the traditional malaria-fighting Chloroquine drug and unresponsiveness is currently expanded to the entire state. The border troops watching the 856 km long India's border with Bangladesh have suffered enormously from malaria being out for more than 15 hours per day outside facing mosquitoes bites.

Malaria and Environment

Malaria has been identified as a major health problem in the northeast India, including Assam, Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, and Tripura states (Mohapatra et. al., 1998). Forested areas and foothills are highly endemic for malaria, caused by P. falciparum with a high prevalence rate. Malaria in Tripura has more than one operating vectors. Anopheles (A) minimus is the major mosquito vector while *A. dirus* also plays a role in transmitting malaria (Prakash et al., 2006) and posing serious health hazards (Das and Baruah 1985; Dutta et el. 1989, Prakash 2005). Deciduous moist forest in Tripura creates favorable conditions for *A. dirus* mosquitoes (Aruna Srivastava et. al.2001).

Tripura in sub-tropical climate with wet, warm and humid weather (P.S. Chaudhuri et al. 2011) has three seasons: summers, winters, and monsoons. Summer continues from the end of March to the end of May (Raja Chakrabortya et al. 2012). In late May, pre-monsoon showers soak the region and temperature soars up to 35°C. Monsoon season arrives normally in June and continues through September (http://www.indian-states.com/tourism.php/Tripura). Heavy downpours are common and average monsoon period rainfall reaches 2,100 mm. The winter season lasts from November to February when the average minimum temperature is around 10°C and rainfall is minimal 12.4 mm in (IMD 2011). Fig. 1. provides some information about precipitation and temperature on one of the Tripura weather stations located in the malarias region (USDA 1994)

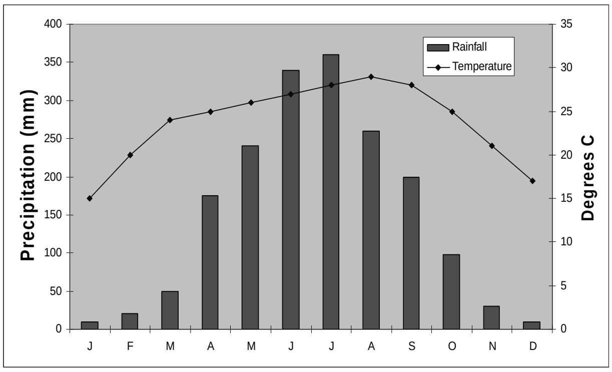

FIG. 1 CLIMATE OF GUAHATI, TRIPURA STATE,INDIA

During rainy season, malaria vector population increases considerably triggering increase in the number of malaria cases during summer months. However, malaria infection reports started in April onwards due to the start of vector population increase. Rains during June through August trigger increase in malaria-positive cases during September to October. From November, vector population declines due to rainfall and temperature decrease during winter season (*Prafull,Dutta* et.el., 2010). (http://www.iloveindia.com/states/tripura/weather.html).

Inter annual variability of malaria incidence follows variations in moisture and thermal conditions. There are casual relationship between malaria and rainfall which creates breading sites for mosquitoes and suitable conditions for longevity through increased relative humidity. However, the relationship with rainfall is often nonlinear because excessive rainfall may wash out mosquitoes immature stages resulting in malaria reduction (Lindsay SW et el. 2000, Najera et. el. 1998).

Data

Malaria incidence and satellite data have been used in this research. Malaria statistics were represented by the annual total clinical malaria cases during 1997-2006. The data were obtained from the India's

Ministry of Health. These data provided percent malaria cases from all patients, who came to the hospitals with fever. The hospital data aggregated to a local administrative unit health centers and to district level included the number of persons tested and found positive for malaria. In this study, the numbers of malaria cases was calculated as slide positivity rate or percent of malaria cases from the total number of victims tested.

Satellite data were presented by weekly Normalized Difference Vegetation Index (NDVI) and Brightness Temperature (BT) collected from the NOAA Global Vegetation Index data set during 1997 to 2006. The Global Vegetation Index (GVI) was developed from the reflectance/emission observed by the Advanced Very High Resolution Radiometers (AVHRR) of NOAA polar-orbiting satellite in the visible (VIS), near infrared (NIR) and infrared (IR) wavelength (Kidwell et. el., 1997). In developing the GVI, the measurements were spatially sampled from 4 km^2 (Global Area Coverage–GAC) to 16 km^2 and from daily observations to seven-day composite(Kidwell et. el. 1997). The VIS and NIR reflectance were pre- and post-launch calibrated and NDVI was calculated as (NIR-VIS)/(NIR+VIS). The IR measurements in 10-11 μm wavelength were converted to BT. Since NDVI and BT have high frequency noise related to variable transparency of the atmosphere, bidirectional reflectance, orbital drift, random and others, which makes it difficult to use, this noise was removed from the data applying statistical techniques to NDVI and BT time series (Kogan 1999). The 1997-2006 weekly NDVI and BT data were collected for each 16 km^2 pixel of the most populated Tripura area, which have the largest percentage of malaria cases. Finally, the average values of NDVI and BT for the highly malarious area (Fig. 2.) for each week and year of the investigated period were calculated and matched with malaria data.

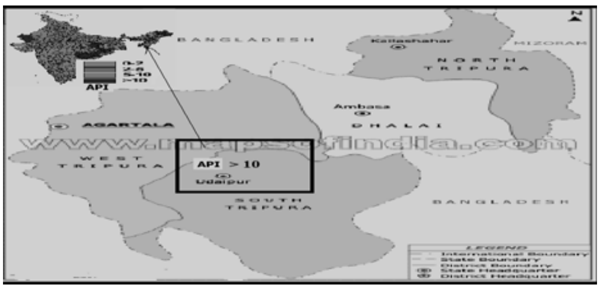

FIG. 2 DISTRICT MAP OF TRIPURA STATE, INDIA RECTANGULAR BOX IS POINTED TO SATELLITE DATA COLLECTION

Methodology

The aim of this research was to develop models for monitoring malaria incidence based on relationship between number of annual malaria cases and VH satellite parameters characterizing weather- and climate-related conditions for each week of the investigated period. In a very general sense, malaria incidence is affected by environmental and socioeconomic/policy factors in addition to biological determinants. Following Obuhov (1949) and Brockwell and Davis (2000), the time series data regulated by both socioeconomic/policy and weather factors should be split into two components deterministic characterizing the first group and random characterizing the second group. In case of malaria in Tripura this can be written as

$$Y = Ysp + dYw \tag{1}$$

Where Y is observed malaria cases, Ysp is malaria cases regulated by socioeconomic and policy factors (deterministic component), dYw is a malaria cases regulated by weather fluctuations (random component).

The Ysp time series are normally expressed by polynomial approximated by the least squares method. If time series are short as in our case then linear polynomial ($Ysp_t = a_0 + a_1 * Y_t$ where a_0 and a_1 are regression coefficients and t is time, in our case year number) is an appropriate approximation of the socioeconomic/policy trend. The dYw is expressed as a deviation from the trend either as a ratio or difference. A ratio ($dYw_t = 100 * Y_t / Ysp_t$) expressed in percent is preferable approximation. As the result, the dYw_t represents deviation (anomaly) of malaria incidents from the trend due to annual and seasonal weather fluctuations.

The principle to construct VH indices stems from the property of green vegetation to reflect/emit solar radiation in the VIS, NIR and IR wavelength. If vegetation is healthy it reflects little radiation in the VIS (due to high chlorophyll absorption of solar radiation), much in the NIR (due to scattering of light by leaf internal tissues and the low absorption by pigments, water and other leaf constituents) resulting in high NDVI. Also healthy vegetation emits less thermal radiation in the IR spectral bands (because more of the available energy is partitioned into the latent heat flux, compared to the sensible heat flux) resulting in relatively cooler canopy and lower BT. Details of the algorithm have been presented in Kogan (1997). We have calculated two indices (equations 2 and 3); Vegetation Condition index (VCI) and Temperature Condition index (TCI). VCI characterizes greenness, vigor and through them the amount of chlorophyll and moisture in vegetation while TCI

characterizes how hot land surface is which includes vegetation canopy and non-vegetated areas.

$$VCI=100*(NDVI–NDVI\ min)/(NDVI\ max–NDVI\ min) \quad (2)$$

$$TCI=100*(BT\ max – BT)/(BT\ max–BT\ min) \quad (3)$$

where NDVI, NDVImax, and NDVImin (BT, BTmax, and BTmin) are no noise weekly NDVI and BT, their multi-year absolute maximum and minimum, respectively. VCI and TCI algorithm separates weather component from ecosystem component presented by the difference between the multi-year absolute max and min NDVI and BT (Kogan et. el. 1995, 1997, 2000). The indices change from 0 to 100 corresponding to variation in weather conditions from extreme stressed to favorable, respectively.

One of the principles used in data analysis was to investigate relationship between malaria data (*in situ)* and satellite data. Therefore, correlation and regression analysis were used in this study. Two types of regression models relating number of malaria cases (dependent variable) against the VH indices (independent variable) were tested with the Ordinary Least Square regression (OLS) and the Principal Component Regression (PCR). Since VH indices (VCI and TCI) values for the neighboring periods were highly correlated (collinear), the PCR method decomposed and transformed the independent variable into a new set of orthogonal, uncorrelated variables (free of collinearity) called principal components of the correlation matrix. This transformation sets the new orthogonal variables in the order based on their importance. The PCR method allows one to estimate regression coefficient more accurately (Draper and Smith 1981, Raymond 1986).

These new regression coefficients are principal component estimators (Gunst and Mason 1980). Finally these coefficients of the newly performed regression analysis were mathematically transformed into a new set of coefficients, which correspond to the original correlated set of variables in the equations.

An important step in building statistical models is to test model predictions independently, since tests using the same training data would be very optimistic. Since the data sample used in this study is limited, a cross–validation technique was applied to independent models tests. In this approach, a single year was left out one-by-one from the data set, a model was built for each set without of the removed year and prediction was made for the eliminated year. As the result, for the 1997-2006 data period, ten independent comparisons between model predictions and observations were made. The models were evaluated based on determination coefficient (R^2), standard error of prediction, student test, bias ((B) - difference between observed (O) and predicted (P) number of malaria cases) and relative bias ((RB)=100*(B)/O) values.

Result and Discussion

Annual Malaria Dynamics

Tripura's malaria annual time series incidence during 1997-2006 are shown in Fig. 3. The Ysp was approximated by the following equation Ysp = - 61 + 0.04 * Year. This estimation indicates that over the investigated period of 10 years, the number of malaria cases in Tripura has slightly increased with the rate of 4% additional cases per year. Although the rate is insignificant, this growth indicates that in spite of the measures taken by the State Government to fight malaria the number of malaria cases is increasing at a stable rate. Fig. 3. also shows that the number of malaria cases fluctuates around the trend. This year-to-year deviation represents a random component (dYw) which depends on weather fluctuation. The dYw expressed as a ratio of the observed to the trend estimated percent malaria cases fluctuates. As seen, the number of malaria cases changes one percent above and below trend level from year to year and the absolute variation reaches two percents, which might account for several hundreds of ill people**.**

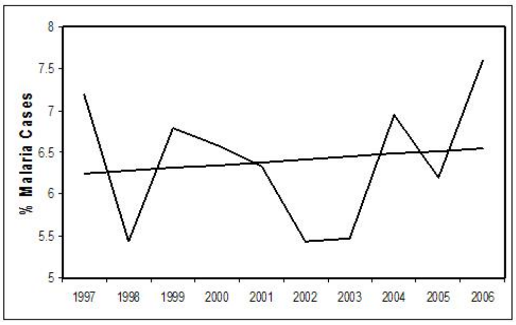

FIG. 3 ANNUAL PERCENT OF MALARIA POSITIVITY DURING 1997-2006, TRIPURA STATE, INDIA

VH Dynamics in the Extreme Malaria Year

To get preliminary understanding if weather conditions expressed through VH indices, have some useful signals about potential for development of malaria, two extremely opposite years in the number of malaria cases (above and below trend) in Tripura have been selected dynamics of VH indices have been compared during these years in the selected malaria prone areas. The extreme malaria year is 2006, with an

elevated malaria cases (above trend) while the year 2002 with least malaria case (below trend). It should be indicated that only two percent difference between the extreme years might be represented by several hundred differences in the number of affected people. Fig. 4. shows dynamics of VCI and TCI for these years. In case of TCI (temperature conditions), two periods during the annual cycle are clearly observed: in spring and fall. The year with more malaria cases (2006) has hotter conditions in spring (weeks 15-19) and cooler in fall (weeks 42-45). When the numbers of malaria cases was smaller (2002) the conditions are opposite: cooler in spring and hotter in fall. The VCI (moisture index) comparison between 2002 (less cases) and 2006 (more cases) in Fig. 4 (b) shows that moisture conditions in summer (weeks 26-36) stimulate mosquitoes' activities including malaria transmission and in fall the situation is opposite – wetter conditions restrain mosquitoes activity.

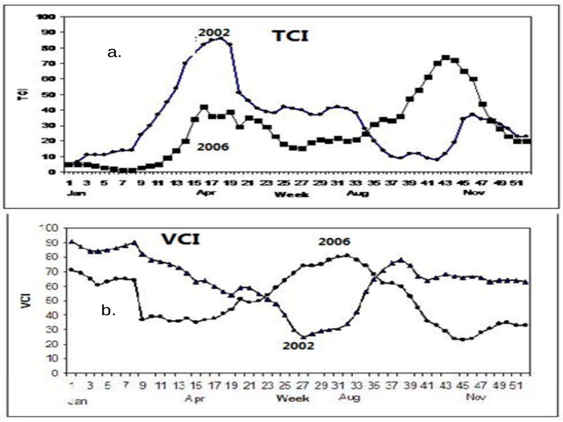

FIG. 4 DYNAMICS OF VEGETATION (VCI) AND TEMPERATURE (TCI) CONDITION INDEX DURING THE TWO YEARS OF EXTREME NUMBER OF MALARIA CASES: LARGE (2006) AND SMALL (2002) IN THE 1997-2006 PERIOD.

Correlation Analysis

Analysis of VH indices dynamics in the extreme year of malaria incidence encouraged us to correlate annual number of malaria cases with weekly VCI and TCI values. Since malaria data set was presented by the total number of malaria cases per year (seasonal and/or monthly data were not available) this level of analysis included correlation of total annual number of malaria cases with VCI and TCI values for each week. The goals were to investigate if (a) dYw correlates strongly with VH indices during the season when weather favors mosquitoes development and their elevated activities (b) dYw correlates weakly during the seasons with non-favorable weather and (c) select the periods when VCI and TCI strongly correlate with the number of malaria cases to be included in the dYw=f(VCI, TCI) models.

Fig. 5. shows correlation coefficient dynamics of total annual number of Tripura's cases (dYw) with weekly VCI and TCI during 1997-2006. As seen, in cooler months (January–February) when mosquitoes are less active due to cold weather; correlation is low for both indices. During the period from spring to early summer (pre monsoon) when mosquitoes activity season begins, correlation starts to increase reaching the maximum of 0.5-0.6 for VCI during June-July (weeks 23-32) and around -0.6 for TCI in weeks 16-19 (April) and later at week 25 (mid June). The relationship between the malaria cases with TCI from Spring to early summer implies that hotter conditions during monsoon period (TCI below 40 indicates an intensification of hot conditions) are accompanied by epidemics increase. In fall (weeks 44-47), dYw versus TCI correlation increases again to nearly 0.6 although it has opposite sign indicating that a very hot temperatures (TCI is below 40) reduce the number of infected people. This conclusion is in line with the results in Figure 4. Summer in Tripura is the central season for monsoon characterized by abundant of rainfall, which creates good conditions for mosquitos' development, activities and spread of malaria. VCI's correlation coefficients (CC) dynamics reflects well such conditions, when CC is gradually increasing from zero values in April towards nearly 0.6 in June and July (weeks 22-32).

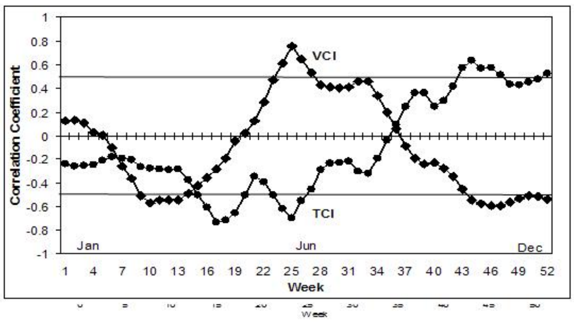

FIG. 5 DYNAMICS OF CORRELATION COEFFICIENTS OF THE TOTAL ANNUAL NUMBER OF MALARIA CASES DEPARTURE FROM THE TREND (dY) WITH WEEKLY VCI AND TCI.

Regression Analysis

Finally, regression models were built and tested following the results shown in Figure 5 and 4. Two statistical approaches OLS and PCR were tested. At the beginning, how OLS model regresses dYw versus independent variables TCI for weeks 17 to 20 weeks has been investigated. The results are presented in Table 1 with the assessment of statistical reliability of

model by t-values, probability of P>|t| and variance inflation assessing collinearity of the independent variables. As seen in the Table, although the R^2 is high (0.87), this model is not reliable, which is indicated by a high values (larger then 0.1) of regression coefficients assessment and a very high values of variance inflation, especially for TCI18 and TCI19. This is confirmed by correlation matrix in Table 2, indicating high CC for the neighboring periods. From this analysis, the following can be concluded. (1)

TABLE 1 OLS MULTIPLE LINEAR REGRESSION OF DY_W VERSUS TCI15 THROUGH TCI20 FOR TRIPURA, R^2=0.87

Variable	Parameter Estimate	t Value	Pr > |t|	Variance Inflation
Intercept	119.1941	10.02	0.0021	0
TCI15	0.8867	1.46	0.2407	25.153
TCI16	-0.52946	-0.55	0.62	56.031
TCI17	-0.68417	-0.77	0.4962	48.346
TCI18	1.01124	0.56	0.6167	167.28
TCI19	-3.32523	-1.92	0.1501	138.32
TCI20	3.68654	2.44	0.0925	45.099

TABLE 2 CORRELATION MATRIX FOR WEEKLY TCI, TRIPURA

	TCI15	TCI16	TCI17	TCI18	TCI19	TCI20
TCI15	1	0.9501	0.8487	0.7537	0.5674	0.3756
TCI16		1	0.9488	0.8697	0.7025	0.5477
TCI17			1	0.9501	0.8238	0.6921
TCI18				1	0.9525	0.8598
TCI19					1	0.9627
TCI20						1

Since the period of observation is too short, the smallest number of independent variables should be selected less than four. (2) Only those variables should be included in the models which have correlation coefficient of dYw with TCI and VCI greater than 0.55. (3) If a few neighboring variables satisfy the condition (2) then mean of a few weeks VCI and/or TCI values should be used. (4) If VCI and TCI show high correlation with dYw for the same weeks of the year then one of the indices should be selected. (5) Addressing collinearity issues, the PCR approach is expected to be a better option although further analysis investigates the independent modeling results for two methods.

1) *Compared of OLS and PCR Models*

Following the indicated criteria, and correlation analysis (Fig 5), four models were selected with the following set of the independent variables: (A) TCI17, TCI25, VCI24; (B) mean TCI for weeks 16-18 (TCI16-18), mean TCI24-26 and mean VCI 23-25; (C) the same three variables as in (A) and TCI43; (D) the same three variables as in (B) and mean TCI42-44.

First, we investigated which of the methods, OLS or PCR, showed better validation results for one of the models, specifically model (D). The independent evaluation of this model for the two methods is shown in Table 3. As seen, PCR-based model when only one principal component was used showed better results because the mean of 10-year independently simulated value is equal to the observed mean but smaller compared to the OLS mean value (6.21%) and the mean bias for PCR is equal to 0 compared to -019 for OLS. Model (D) was also tested for the number of principal component (PC) elements (Table 4). Based on *t* values and Pr>*t* it is possible to conclude that only the first PC is statistically significant with the 97% probability for the two and 94% for the three PC. In summary, the first PC is sufficient for PCR modeling.

TABLE 3 OBSERVED AND INDEPENDENTLY SIMULATED PERCENT MALARIA CASES FOR MODEL (D) USING PCR AND OLS MODELS, TRIPURA, INDIA

Year	% Cases Observed	Simulated OLS	Simulated PCR	BIAS OLS	BIAS PCR
1997	7.19	6.44	6.96	-0.75	-0.23
1998	5.44	5.80	5.34	0.36	-0.10
1999	6.79	6.29	6.22	-0.50	-0.57
2000	6.58	6.19	5.90	-0.39	-0.68
2001	6.33	6.44	6.55	0.11	0.22
2002	5.43	4.19	5.75	-1.24	0.32
2003	5.47	6.04	6.00	0.57	0.53
2004	6.95	7.30	7.26	0.35	0.31
2005	6.2	6.78	6.69	0.58	0.49
2006	7.6	6.66	7.35	-0.94	-0.25
Mean	**6.40**	**6.21**	**6.40**	**-0.19**	**0.00**
Standard Deviation	**0.77**	**0.82**	**0.67**	**0.66**	**0.43**

TABLE 4 STATISTICAL ANALYSIS ON THE NUMBER OF COMPONENTS IN PCR MODEL (D)

R-square = .76

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	100.002	2.09295	47.78	<.0001
Prin1	-6.55308	1.47049	-4.46	0.0029
Prin2	3.39257	2.20754	1.54	0.1682

R-square = 0.77

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	100.002	2.24394	44.57	<.0001
Prin1	-6.55308	1.57657	-4.16	0.006
Prin2	3.39257	2.3668	1.43	0.2017
Prin3	0.96543	3.22382	0.3	0.7747

2) *PCR Analysis*

The PCR method was finally applied to the four indicated above models with the different sets of VH independent parameters. The goals were to investigate if (1) single week VH variables provide

better modeling results compared to the several weeks mean; (2) both moisture (VCI) and thermal (TCI) variables are better predictors compared to either of them; (3) model performance improves after adding VH variables at the end of the year. The equations of four PCR models (A-D) when the first principal component only is used are written below.

(A) $DY = 99.23 - 0.094\ TCI_{17} - 0.72\ TCI_{25} + 0.25VCI_{24}$; $R^2 = 0.45$

(B) $DY = 105.64 - 0.18TCI_{mean16\text{-}18} + 0.29\ TCI_{mean24\text{-}26} + 0.30\ VCI_{mean23\text{-}25}$; $R^2 = 0.64$

(C) $DY = 97.06 - 0.096\ TCI_{17} + 0.32\ TCI_{25} + 0.27\ VCI_{24} + 0.05\ TCI_{43}$; $R^2 = 0.54$

(D) $DY = 103.65 - 0.19\ TCI_{mean16\text{-}18} - 0.31\ TCI_{mean24\text{-}26} + 0.30\ VCI_{mean23\text{-}25} + 0.06\ TCI_{mean42\text{-}44}$; $R^2 = 0.69$

These results as shown in table 5 answered the three questions set above. The R^2 values for the entire regression indicate that the equations with single week VH variables are less reliable compared to several weeks mean since they have smaller R^2 (model (A) $R^2 = 0.45$ is significantly smaller than model (B) $R^2 = 0.64$ and model (C) $R^2 = 0.54$ is significantly smaller than (D) $R^2 = 0.69$). Using PCR approach both moisture (VCI) and thermal (TCI) variable can be used in the equation. Adding variable at the end of malaria season improves regression estimates. For an early prediction of malaria in Tripura the equation (B) could be used, which will provide estimates at the end of June.

TABLE 5 PCR MODELS A,B,C AND D,TRIPURA, INDIA

A.	R-Square =	0.45		
Variable	Estimate	Error	t Value	Pr > |t|
Intercept	100.00	3.00	33.28	<.0001
Prin1	-5.70	2.29	-2.49	0.0378
B.	R-Square =	0.64		
	Parameter	Standard		
Variable	Estimate	Error	t Value	Pr > |t|
Intercept	100.00	2.54	39.39	<.0001
Prin1	-6.22	1.81	-3.44	0.0088
C.	R-Square =	0.54		
	Parameter	Standard		
Variable	Estimate	Error	t Value	Pr > |t|
Intercept	100.00	2.75	36.39	<.0001
Prin1	-6.21	2.08	-2.99	0.0173
D.	R-Square =	0.69		
	Parameter	Standard		
Variable	Estimate	Error	t Value	Pr > |t|
Intercept	100.00	2.26	44.17	<.0001
Prin1	-6.55	1.59	-4.12	0.0033

Independent Validation

Independent validation results using "one-out-one-in" method are shown in Fig 6 and Table 3. PCR model with the one PC element shows better results because its 10-year mean value (6.40%) is equal to the observed mean (6.40%) compared to a lower OLS mean value (6.21%), the mean bias for PCR is equal to 0 compared to -019 for OLS, and R^2 for PCR is much higher 0.68 versus 0.43 for OLS. As shown in Fig 6, the OLS-simulated value for 2002 has very large bias (-1.24). But even if this observation is removed from the OLS regression analysis R^2 fails to get larger.

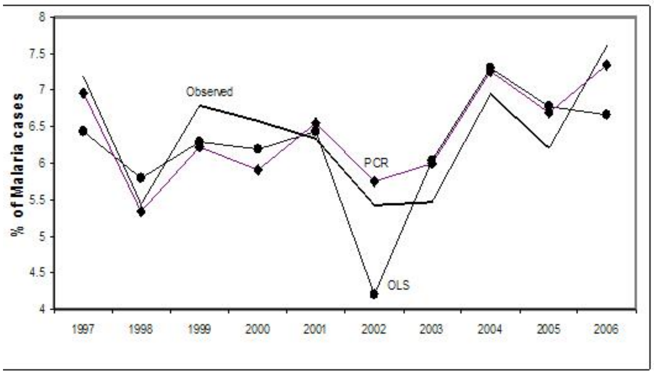

FIG. 6 TIME SERIES OF OBSERVED AND INDEPENDENTLY SIMULATED PERCENT MALARIA CASES USING TWO REGRESSION METHODS (PCR AND OLS)

Conclusions

Tripura is considered as one of the malaria endemic areas in India. This research has shown that malaria transmission in Tripura state is sensitive to both thermal and moisture conditions, which can be estimated from NOAA/AVHRR-based Vegetation health indices data. The most sensitive periods of malaria to the environment are the end of March-early April (weeks 14-17) using TCI, second half of June (weeks 23-26), using both TCI and VCI and mid October (weeks 42-44) using TCI. These parameters can serve as independent variables for prediction of malaria epidemics in Tripura, India. It is important to indicate also that model (B) can be also used for an early detection of malaria because the results of independent evaluation of this model are as good as the model (D). Meanwhile, the model (B) taking some advantages since the prediction is given in June can be instrumental to predict malaria at least three months earlier than the starting of main transmission high season (post rainy season August and September). Finally, it should be mentioned that VH indices derived from AVHRR sensor on NOAA polar orbiting satellites are delivered in real time (every Monday) to the NOAA/NESDIS web site (http://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/vh_browse.php) providing valuable information about malaria epidemics around the world including Tripura in India in order to take preventive measures in due time.

ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers for their valuable comments which significantly improved the quality of the paper.

REFERENCES

Aruna Srivastava, B. N. Negpal,Rekha Saxena and S.K. Subbarao,(2001) Predictive habitat modeling for forest malaria vector species *An. Dirus in* India – A GIS – based approach, Current Science , Vol 80, No.9

Brockwell, P. J. and Davis, R. A., 2000, Introduction to time series and forecasting (New York: Springer).

Das SC, Baruah I ,(1985), Incrimination of Anopheles minimus Theobald and Anopheles balabacensis balabacensis Baisas (A. dirus) as malaria vectors in Mizoram. Indian Journal of Malariology 22:53

Dutta P, Bhattacharyya DR, Sharma CK, Dutta LP (1989) The importance of Anopheles dirus (A. balabacensis) as a vector of malaria in northeast India. Indian Journal Malariology 26:95

Dreaper, N.R. And Smith, H., 1981, Applied Regression Analysis (New York: Wiley).

Gunst, R.F. and Mason, R.L., 1980, Regression Analysis and its Application: a Data oriented Approach (New York: m. Dekker).

IMD (India Meteorological Department) 2011: http://www. imd.gov.in/section/hydro/distrainfall/webrain/tripura/northtripura.txt

Kidwell, K. B., 1997, Global Vegetation Index user's guide (Camp Springs MD: US Department of Commerce, NOAA, National Environmental Satellite Data and Information Service, National Climatic Data Center, Satellite Data Services Division).

Kogan, F.N.,1995, Droughts of the late 1980s in the United State as derived from NOAA polar orbiting satellite data, Bulletin of the American Meteorological Society ,76:655-668.

Kogan, F.N., 1997, Global drought watch from space, Bulletin of the American Meteorological Society,78, 621-636.

Kogan, F.N., 2000, Global drought detection and impact assessment from space. In Drought: A Global Assessment (D.A. Wilhite,Ed) V0l,I, Hazard and Disaster Series, Routledge, London and New York, 196-210.

Lindsay SW, Bodker R, Malima R, Msangeni HA, Kisinza W, 2000, Effect of 1997-1998 El Nino on highlan malaria in Tanjania, Lancet 355: 989-990

Mohapatra PK, Prakash A, Bhattacharyya DR, Mahanta J (1998) Malaria situation in north-eastern region of India. ICMR Bulletin 28(3):21–30

Najera JA, kouznetzsov R, L,, Delacollette,C., 1998. Malaria Epidemic: Detection and control,Forecasting and Prevention. Geneva: World Health Organization. WHO/MAL/98.1084

Prakash A, Walton, Bhattacharyya DR, Samantha OL, Mohapatra Pk, Mahanta J (2006), Molecular characterization and species identification of the Anopheles dirus and An. minimus complexes in northeast India using r-DNA ITS-2. Acta Tropica 100:156–161

Prakash A, Bhattacharyya DR, Mohapatra Pk, Mahanta J (2005) Malaria transmission risk by the mosquito Anopheles baimai (formerly known as An. dirus species D) at different hours of the night in Northeast India. Medical and Veterinary Entomology 19:423–427

P.S. Chaudhuri & Subhalaxmi Bhattacharjee, (2011), Reproductive biology of eight tropical earthworm species of rubber plantations in Tripura, India, *Tropical Ecology* 52(1): 49-60

Prafulla Dutta, Siraj Ahmed Khan, Dibya Ranjan Bhattarcharyya, Abdul Mabood Khan, Chandra Kanta Sharma, and Jagdish Mahanta , (2010), Breeding habitats of Anopheles baimai and its role in incidence of malaria in Northeastern region of India EcoHealth, DOI: 10.1007/s10393-010-0337-7

Raja Chakrabortya, Biplab Deb, N Devannac, Saikat Sena, (2012), North-East India an Ethnic Storehouse of Unexplored Medicinal Plants, *J. Nat. Prod. Plant Resour*. 2 (1):143-152

USDA, 1994: Major World Crop Areas and Climate Profiles. United States Department of Agriculture, Food and Agricultural Outlook Board. Agricultural Handbook No.664, Washington D.C.

an, A., F. Kogan, L. Roytman, M. Goldberg and W. Guo, 2011**.** Modeling and prediction of malaria vector distribution in Bangladesh from remote sensing data. *International Journal of Remote Sensing* v. 32(5) pp. 1233-1251. (DOI:

080/01431160903527447).

Rahman, A., L. Roytman, M. Goldberg and F, Kogan, 2011: Comparative Analysis on Applicability of Satellite and Meteorological Data for Prediction of Malaria in Endemic Area in Bangladesh. *Am. J. Trop. Med. Hyg.,* 82(6), pp. 1004–1009.

Rahman, A., F. Kogan and L. Roytman, 2006: Short Report: Analysis of Malaria Cases Assessing Bangladesh with Remote Sensing Data. *Am. J. Trop. Med. Hyg.,* 74(1), pp. 17–19.

Raymond H. Myers, 1986, Classical and Modern Regression with Applications, Dexbury Press, Boston Massachusetts

Shiv Lal, G.S. Sonal, P.K. Phukan, 2010; Status of Malaria in India. *Journal of Indian Academy of Clinical Medicine _ Vol. 5 _ No. 1.*

Thomson, M. C., Doblas-Reyes, F. J., Mason S. J., Hagedorn, R. S., Connor, J., Phindela, T., Morse, A. P., and Palmer, T. N., 2006, Malaria early warnings based on seasonal climate forecasts from multi-model ensembles, Nature, 439, 576-579.

Wikipedia,2010:http://en.wikipedia.org/wiki/Tripura#Geography_and_climate

IMD (India Meteorological Department) 2011: http://www.imd.gov.in/section/hydro/distrainfall/webrain/tripura/northtripura.txt

Mohammad Nizamuddin received his Bachelor in Electrical engineering from Bangladesh university of Engineering and Technology (BUET), Masters and Ph.D from the City University of New York, USA. His major field of studies include optical remote sensing; algorithm development for ecosystems, climate and weather impact assessments on vector borne diseases, food and agricultural.

He has been teaching power system analysis, electric circuit analysis, clock design at City University of New York and Laguardia community college for around seven years as a visiting faculty. He is now working as a research associate at the Research Foundation, City University of New York where he was involved in inter-calibration and scaling-up Algorithm for SEVIRI and NOAA-18 data (development, validation and application) project awarded by National Environmental Satellite Data and Information Service Program Office (NESDISPO),International Disaster Management Sensor Web-African Sub-Sahara Flood/Disease Sensor Web Pilot Project funded by NOAA and NASA and United Nations.

Dr. Nizamuddin was a member of IEEE, SPIE, and AMS . He has a good number research outputs published in peer reviewed internationally reputed journals and proceedings to his credit.

Felix Kogan received a M.S. in Agricultural Meteorology from Odessa Meteorological Institute, USSR and a Ph.D. in Environmental Sciences from the World Meteorological Center, Moscow, USSR. After serving as a Senior Scientist at the World Meteorological Center in Moscow, USSR until 1980, he spent five years at the University of Missouri-Columbia as a Research Associate. Felix joined NESDIS in 1985, applying his expertise to a number of areas, including remote sensing; ecosystems; climate and weather impact assessments. Felix has acted as Principal Investigator of numerous national and international projects funded by the Agency for International Development, the United Nations Food and Agricultural Organization, NATO and NOAA. He has lead global-scale international cooperation with government institutions, universities, and private organizations in transferring new technologies. He is a member of the UN World Meteorological Organization (WMO) Technical Team and of the editorial board of the Indian Journal of Remote Sensing.

Ramesh Dhiman currently serves as Scientist F /Deputy Director (Sr Grade) at the National Institute of Malaria Research (Indian Council of Medical Research) under the Department of Health Research, in New Delhi, India. He earned his PhD in Parasitology/Entomology in 1983. Before entering this position, he served at RMRIMS in Patna for work on visceral leishmanisis and National Institute of Virology in Pune. Dr Dhiman is a Co-Leader of the "Group on Earth Observations Health and Environment Community of Practice" and has contributed to work on climate change and health, including efforts with the WHO, IPCC, and Indian government, and is widely published. His fields of interest are epidemiology of malaria including application of satellite remote sensing, climate change & Health, developing tool for early warning of malaria and visceral leishmaniasis and dengue.